

# The Ides of March Is Fast Approaching; Take Heed of Any Warnings in Your Enterprise Data

The Ides of March is fast approaching. Take heed of any warnings in your enterprise data so you can prepare in advance before a hidden risk gets the jump on you. While I don't know exactly what lurks in your data, there's a good chance something represents a hidden risk. Maybe it's a deal that went south. Maybe it's a former client. Maybe it's a former employee. But you won't fully know what's in your data until you deploy enterprise search.

## **What does enterprise search do?**

Enterprise search enables instant concurrent searching through terabytes from a standard Windows network, from an "on premises" web server or from a cloud server. To instantly concurrently search terabytes, enterprise search first needs to index the data. Indexing identifies each individual word and number in the data and its location in the full-text and metadata. A single dtSearch index can hold up to a terabyte of text and there is no limit on the number of indexes the software can create and simultaneously search.

## **How much work is it to start indexing?**

All you need to do is check off the folders to index. dtSearch's "off-the-shelf" Windows software can cover both local files as well as remote files like SharePoint, Office365 and DropBox files that present as part of the Windows folder system. No need to tell the indexer whether a particular file is local or remote or what type of file it is. While the indexer needs to identify the file type of each file before parsing it, the indexer can on its own figure out if each file is a word processing

Article contributed  
by dtSearch®



**Enterprise search  
enables instant  
concurrent searching  
through terabytes  
from a standard  
Windows network,  
from an "on premises"  
web server or from a  
cloud server.**

document, a spreadsheet, a database, a presentation file, a note file, a PDF, an email with attachments, etc.

### **How does the indexer make this determination?**

The indexer uses the binary format for file format identification and subsequent parsing. It doesn't matter if the file has a mismatched file extension like a OneNote file saved with a .PDF extension or a PDF saved with a .ONE extension. The indexer can even process recursively nested files like an email with a ZIP or RAR attachment holding a spreadsheet with a Word document nested inside. By going directly to the binary format and bypassing the retrieval of each file in its associated application, the indexer has the advantage of not only greater speed but also enhanced access.

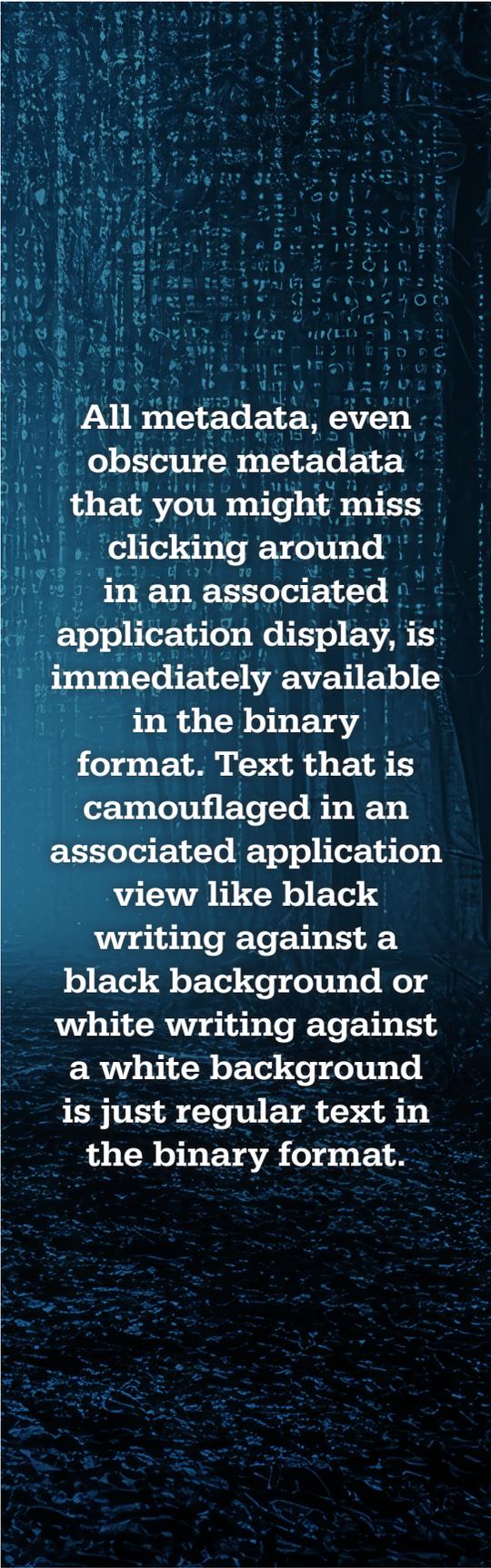
### **Like what?**

All metadata, even obscure metadata that you might miss clicking around in an associated application display, is immediately available in the binary format. Text that is camouflaged in an associated application view like black writing against a black background or white writing against a white background is just regular text in the binary format. The indexer can also access tracked changes that have not been fully accepted, or text marked for redaction that even if not visible by default from an application display still remains with the file. The indexer can even flag "image only" PDFs requiring OCR for full-text indexing.

### **How does dtSearch handle data updates?**

dtSearch can automate index updates so the indexer can tackle only the new, modified or deleted files, without re-doing the entire dataset "from scratch." Critically for enterprise use, re-indexing does not prevent searching, so instant concurrent searching can continue while indexes automatically update.

Article contributed  
by [dtSearch®](#)



**All metadata, even obscure metadata that you might miss clicking around in an associated application display, is immediately available in the binary format. Text that is camouflaged in an associated application view like black writing against a black background or white writing against a white background is just regular text in the binary format.**

## **So what types of searches can you do to identify suspect data?**

Say you were searching for the *ides of march*. You could do an exact phrase search for *ides of march*. Or you could enter an “any words” search finding files that reference either *ides* or *march*. Or you could perform an “all words” search retrieving only files mentioning both *ides* and *march*. You could also add other Boolean and/or/not components to your search, like requiring both *deal that went south* and *former employee* with no mention of *sunny day*.

You could further enter a proximity search for *ides* appearing within 15 words of *march*. Or you could narrow that search request, retrieving only files that contain *ides* within 6 words before *march*. Or you could require *ides of march* in specific metadata. Or you could look for a number or numeric range. Or you could add on a date or date range element in certain metadata or across all text. Searching *3/5/26* to *4/1/26* would pick up both *3/15/26* and *March 15, 2026* in the text.

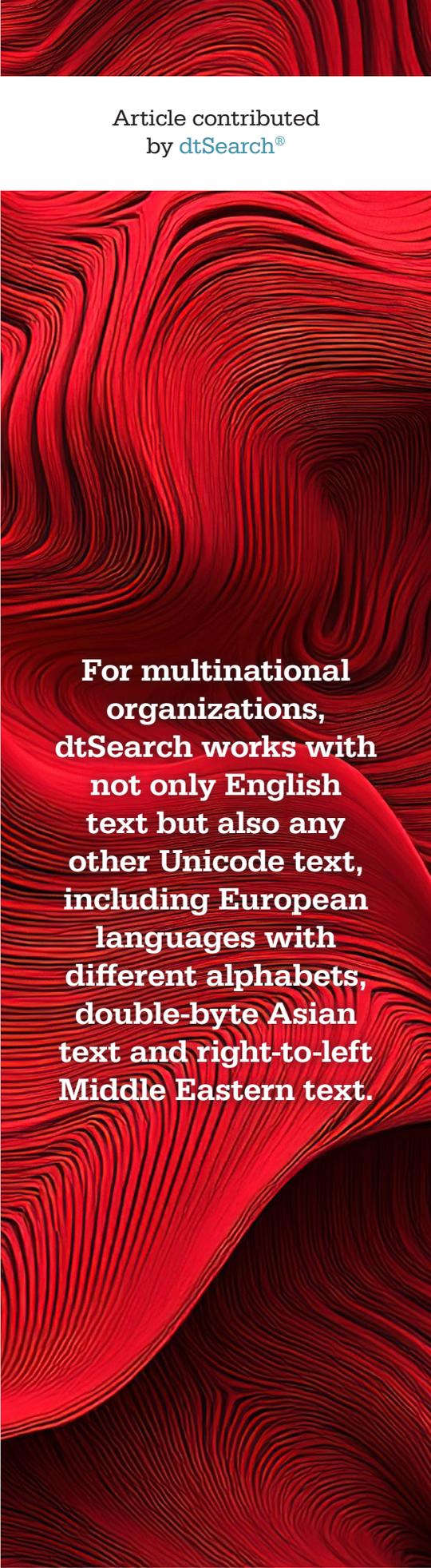
### **Anything else?**

Concept searching lets you enter a synonym like *Joe* for *Joseph*. Fuzzy searching adjusts from 1 to 10 to sift through typographical or OCR errors like *employee* for *employee*. For multinational organizations, dtSearch works with not only English text but also any other Unicode text, including European languages with different alphabets, double-byte Asian text and right-to-left Middle Eastern text. In addition to word-based searching, dtSearch can also search for other Unicode characters like specific Unicode emojis. The product can even flag any credit card numbers that may appear in indexed data.

### **What about sorting?**

Default relevancy ranking will sort by hit density and rarity. Take a search for *ides* or *march*. If *march* is in millions of files while *ides* is in just a few, then *ides* will get a higher relevancy score with files with the densest mentions coming out on top. Users can also define

Article contributed  
by dtSearch®



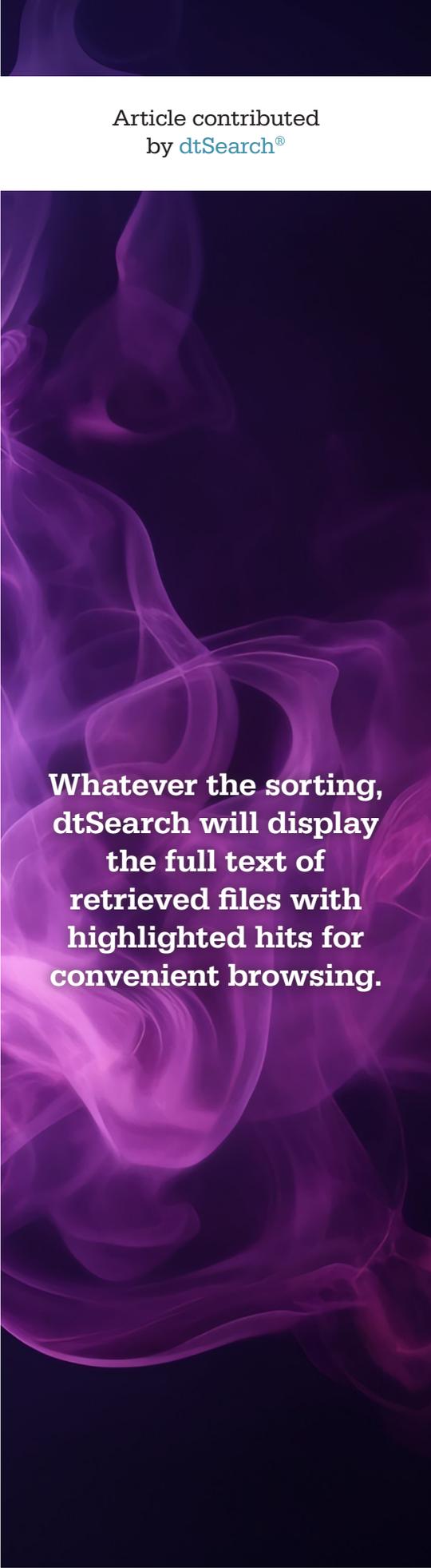
**For multinational organizations, dtSearch works with not only English text but also any other Unicode text, including European languages with different alphabets, double-byte Asian text and right-to-left Middle Eastern text.**

their own relevancy-ranking like giving *ides* a positive weight of 8 and *Caesar* a negative weight of 6. For a different perspective on search results, instantly re-sort search results by a completely different metric like file date or file location. Whatever the sorting, dtSearch will display the full text of retrieved files with highlighted hits for convenient browsing.

### **Final thoughts?**

dtSearch.com has fully-functional 30-day evaluation downloads to start your organization now on instant concurrent searching in anticipation of the Ides of March – or simply to find what you need to get on with your day.

Article contributed  
by [dtSearch®](#)



**Whatever the sorting,  
dtSearch will display  
the full text of  
retrieved files with  
highlighted hits for  
convenient browsing.**